

CHUM: A Frame Supplementation Procedure for Address-Based Sampling

Joe McMichael, Bonnie Shook-Sa,
Jamie Ridenhour, Rachel Harter

Federal Committee on Statistical Methodology Research Conference
Washington DC
November 4, 2013

Contact

Joe McMichael

Research Statistician

919.485.5519

McMichael@rti.org

Overview

- Background
- Intro to CHUM (a general walk through)
- CHUM Summary
- Operational Considerations
- Drawbacks and Benefits
- Case Studies
- Current and Future Research

Background

What do we mean by coverage?

- How well ABS frames - specifically those based on the Computerized Delivery Sequence file (CDS) – represent the U.S. household population.
- Given in the context of area probability, in-person surveys.
- ABS frames are completely or partially replacing household field enumeration in an effort to reduce cost and in some cases improve coverage.

Sources of undercoverage ABS:

- Unlocatable Mailing Addresses (e.g. PO Boxes)
- Geocoding Error
- Frame Latency

National household coverage estimates:

~98% in urban areas.

~77% in rural areas.

Background (continued)

Concern over coverage bias.

- Influence of coverage bias will be vary by outcome and study.
- Nationally speaking, it seems reasonable to assume household under-coverage is not random. Creates the potential for bias. In evidence the urban/rural coverage disparity.
- Survey practitioners are prudent in developing methods to manage coverage bias.

Frame supplementation is not new to in-person household surveys.

- The Half-Open Interval (HOI)

Note: Frame supplementation methods are just one tool that attempt to increase data quality for a given cost. Specifically they attempt to reduce bias introduced by the frame. More work is needed to understand the cost components of these methods and how they relate to controlling the bias and variability of estimates so survey practitioners can choose the methods appropriate for their survey.

The CHUM Procedure

CHUM = Check for Housing Units Missed

In-field protocol that systematically provides a known probability of selection to every dwelling unit not on the ABS frame.

CHUM walk though

Two Components:

1. CHUM 1 - Check for missed units.
 - Based on a sample of addresses
2. CHUM 2 - Check for missed blocks.
 - Based on a sample of census blocks

Basic Path of Travel Rule:

- Move clockwise around the block and do not cross streets.

Special Situations:

- Multi-unit structures (e.g. apartments)
- Other less common situations (e.g. invisible boundaries, missing streets)



ENGLEWOOD AVE

ENGLEWOOD AVE

ENGLEWOOD AVE

ARBOR ST

WASHINGTON ST

EDGEVALE RD

GLENDALE AVE

EVERETT PL

EVERETT PL

EVERETT PL

ARBOR ST

WASHINGTON ST

EDGEVALE RD

GLENDALE AVE

W KNOX ST

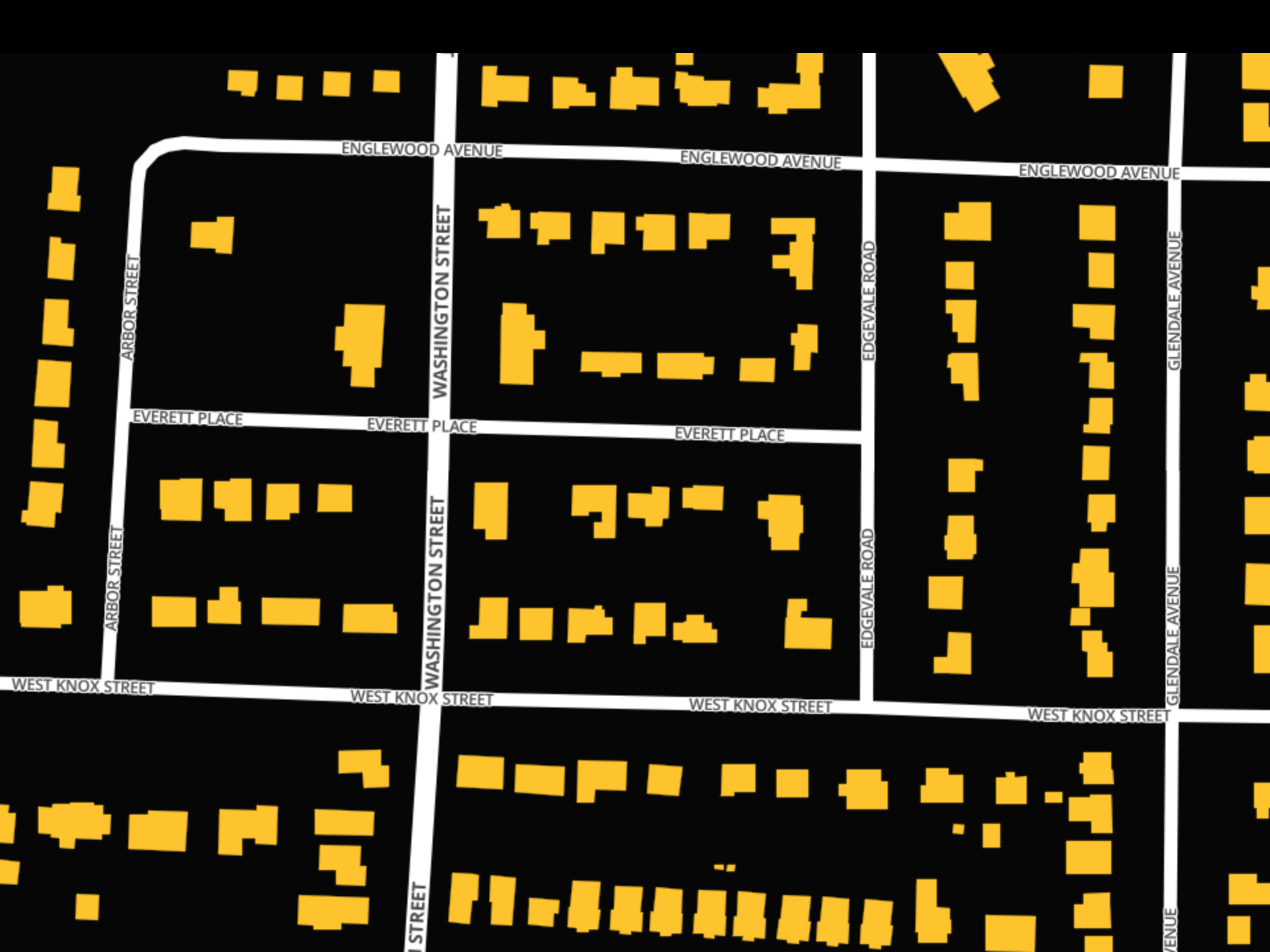
W KNOX ST

W KNOX ST

W KNOX ST

DN ST

AVE



ENGLEWOOD AVENUE

ENGLEWOOD AVENUE

ENGLEWOOD AVENUE

ARBOR STREET

WASHINGTON STREET

EDGEVALE ROAD

GLENDALE AVENUE

EVERETT PLACE

EVERETT PLACE

EVERETT PLACE

ARBOR STREET

WASHINGTON STREET

EDGEVALE ROAD

GLENDALE AVENUE

WEST KNOX STREET

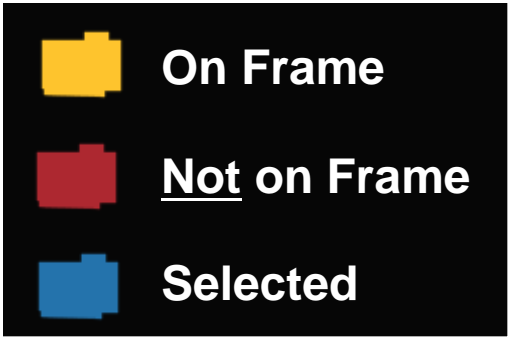
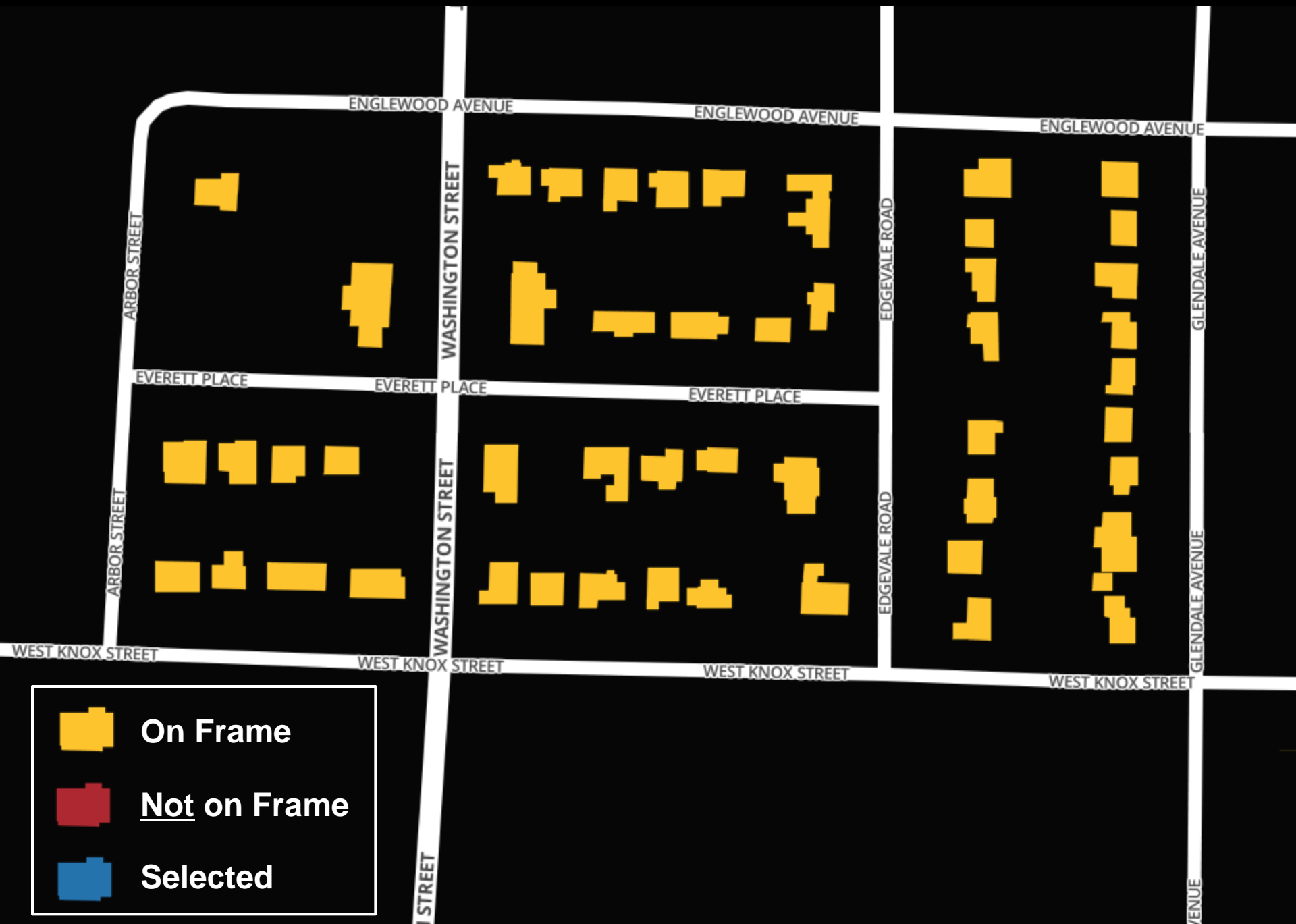
WEST KNOX STREET

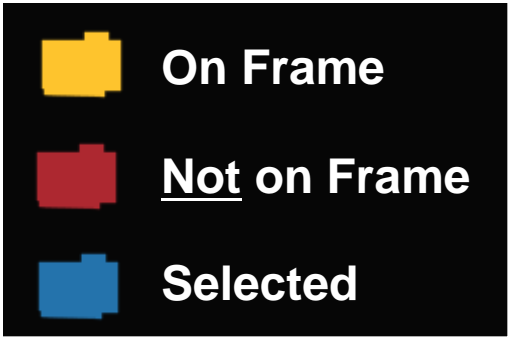
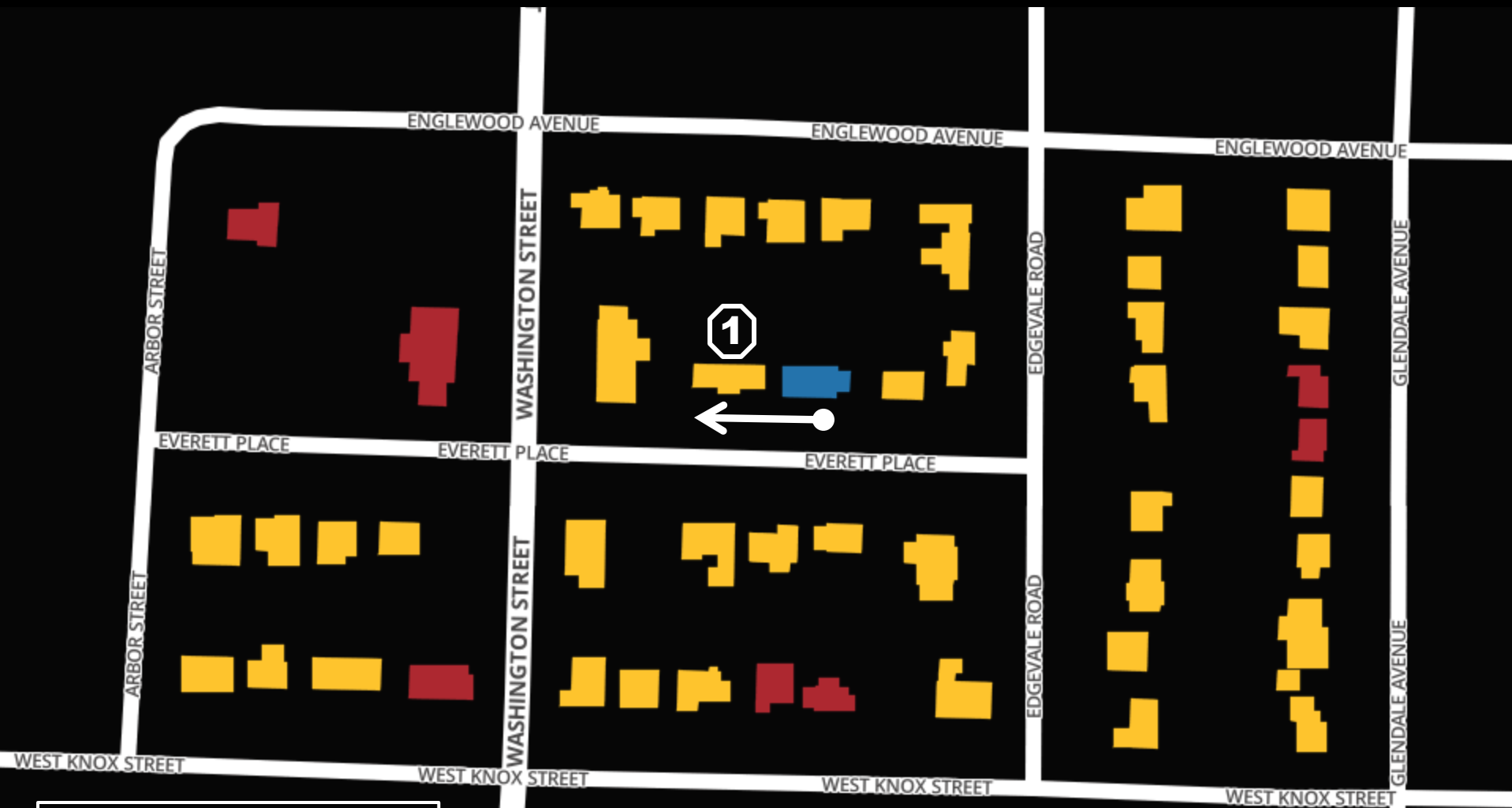
WEST KNOX STREET

WEST KNOX STREET

I STREET

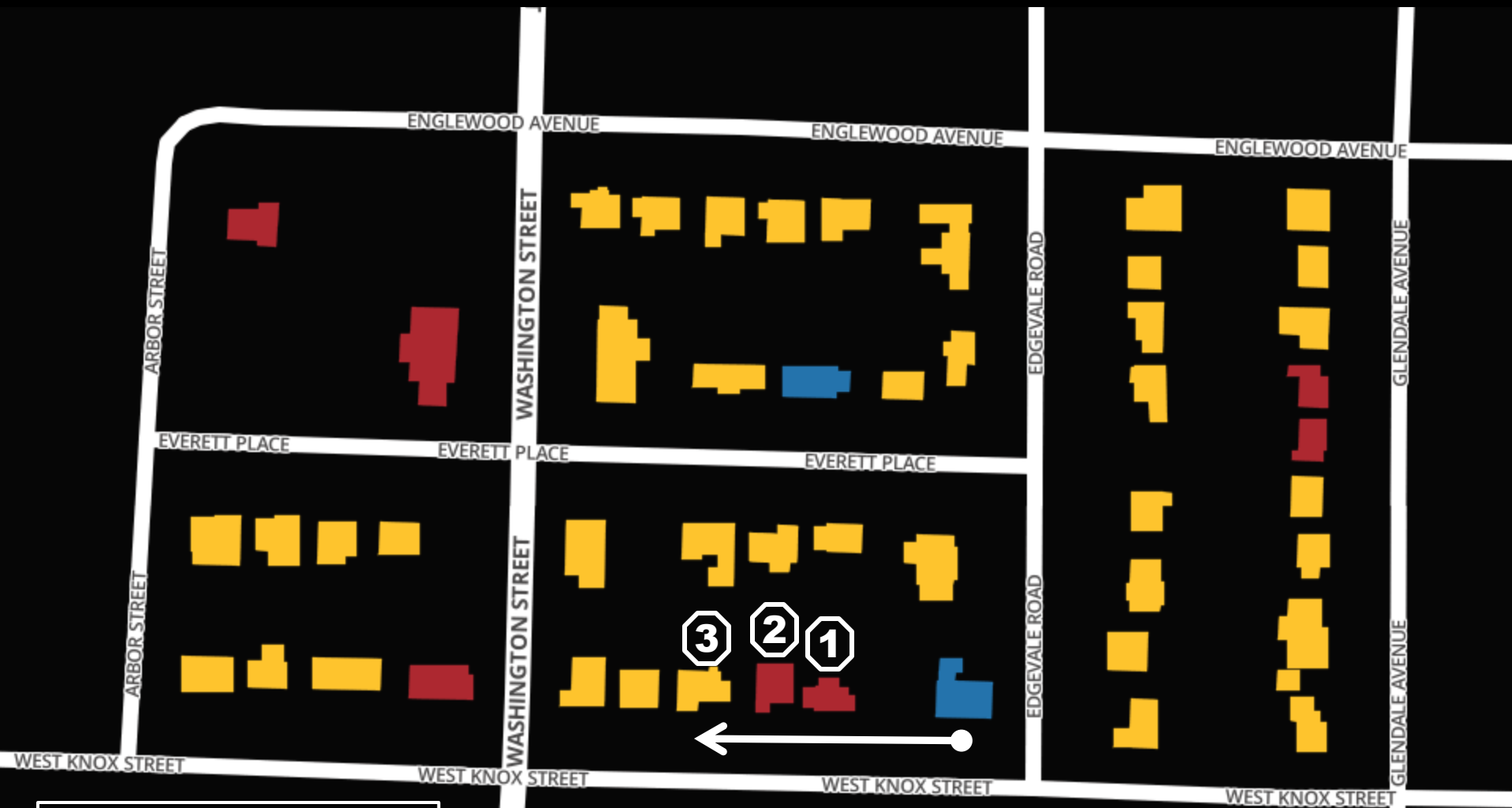
AVENUE







Two components:

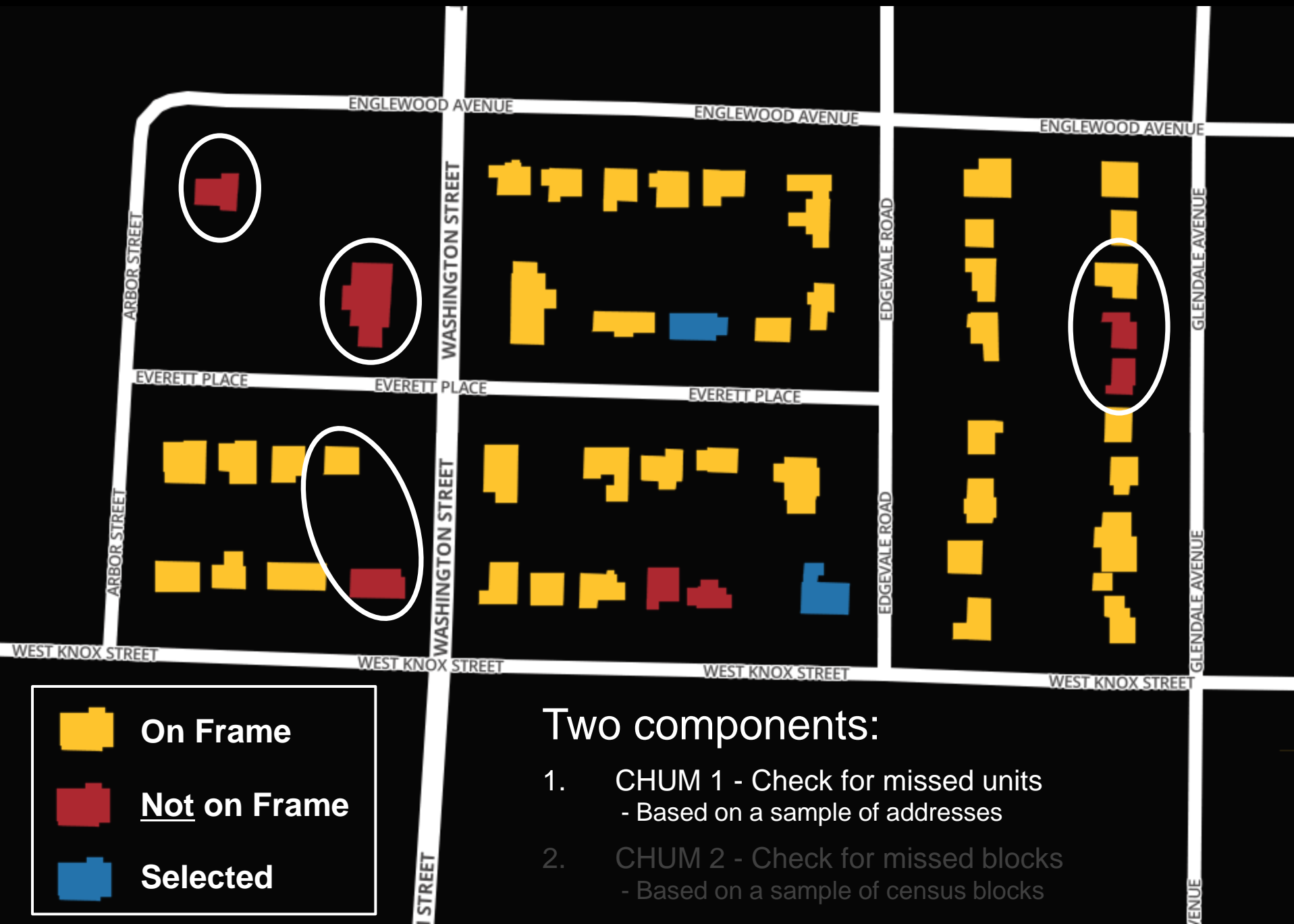
1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks






	On Frame
	<u>Not</u> on Frame
	Selected

Two components:

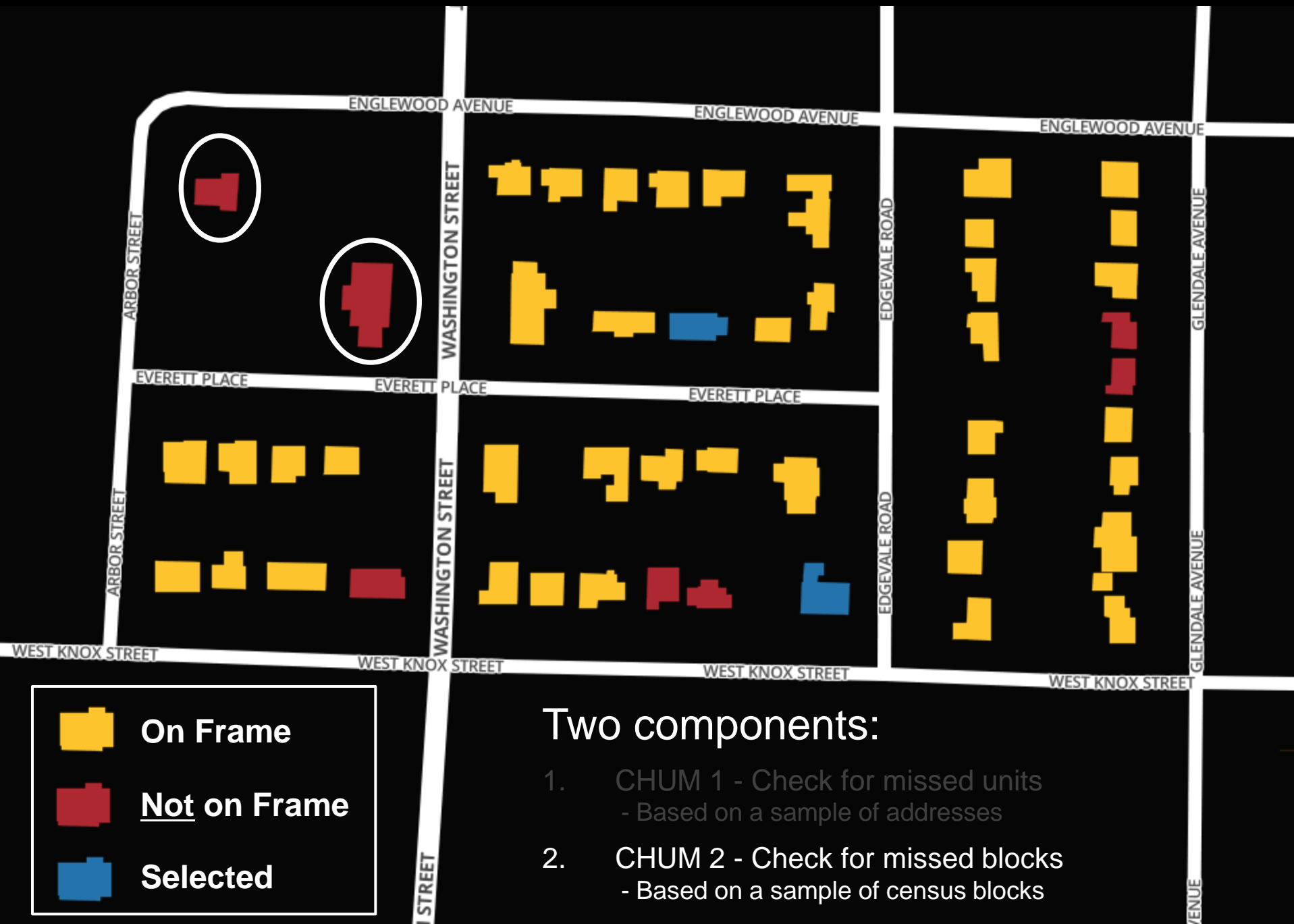
1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks






	On Frame
	<u>Not</u> on Frame
	Selected

Two components:

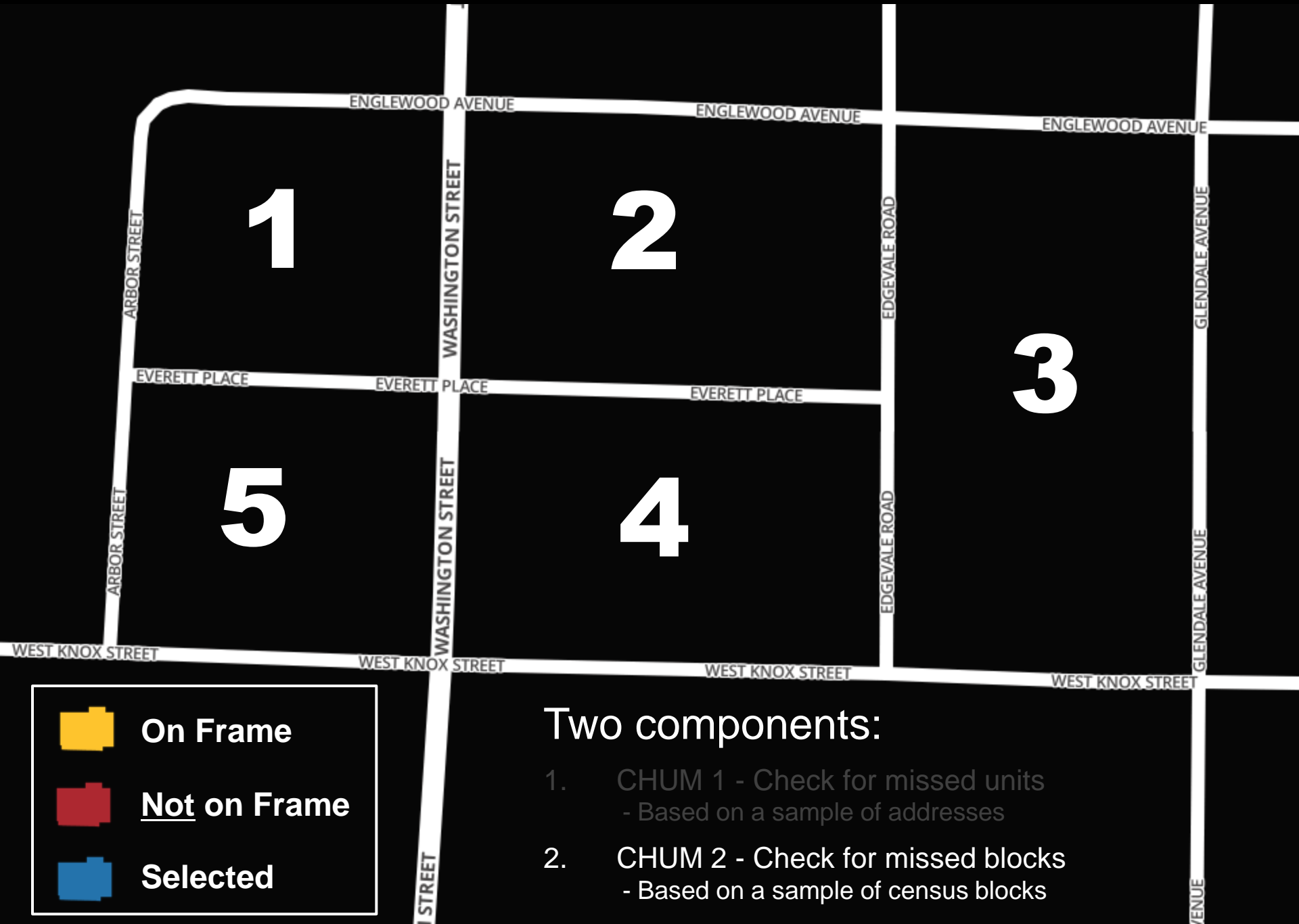
1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks

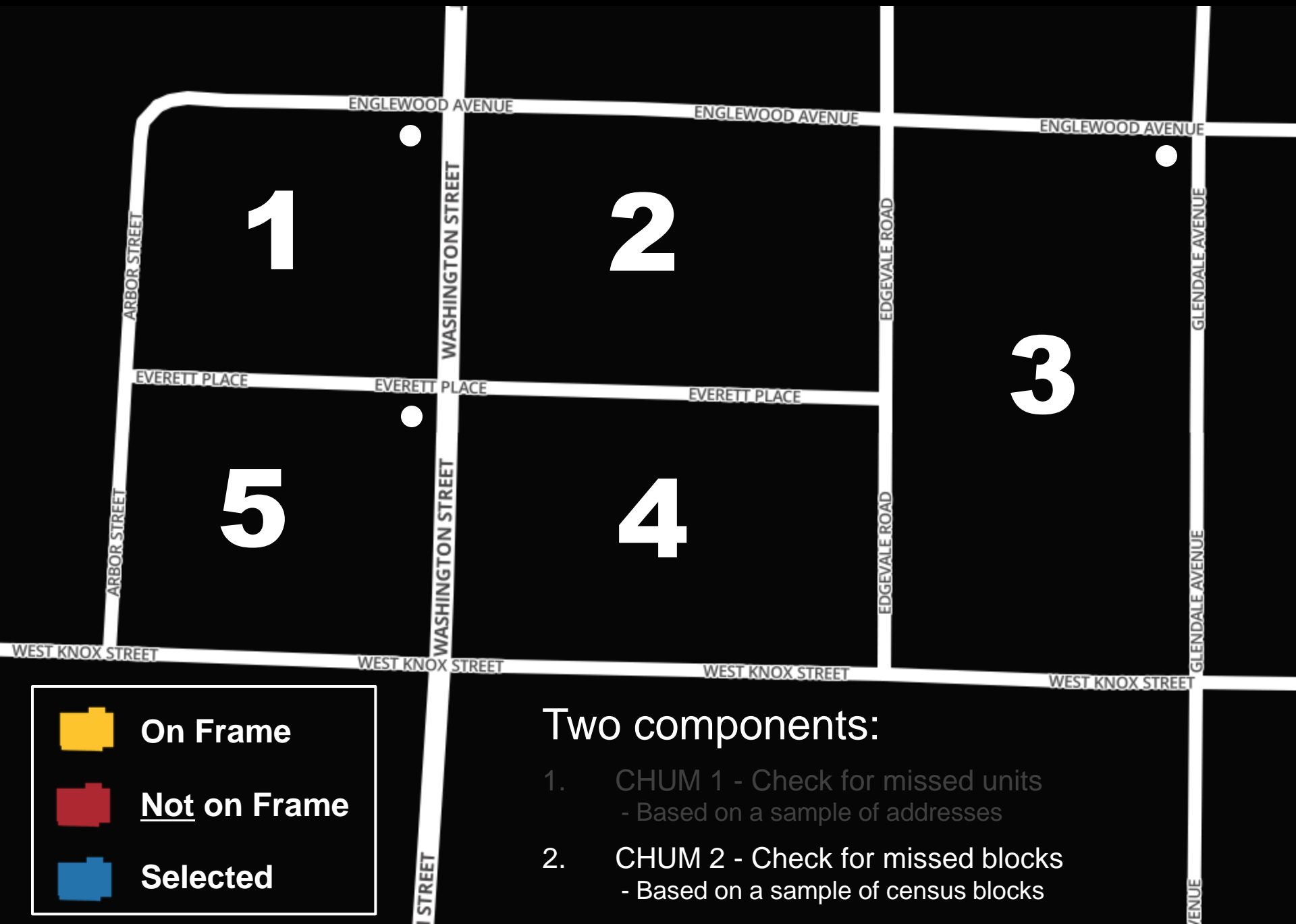


-  On Frame
-  Not on Frame
-  Selected

Two components:

1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks






1


2


3

5

4

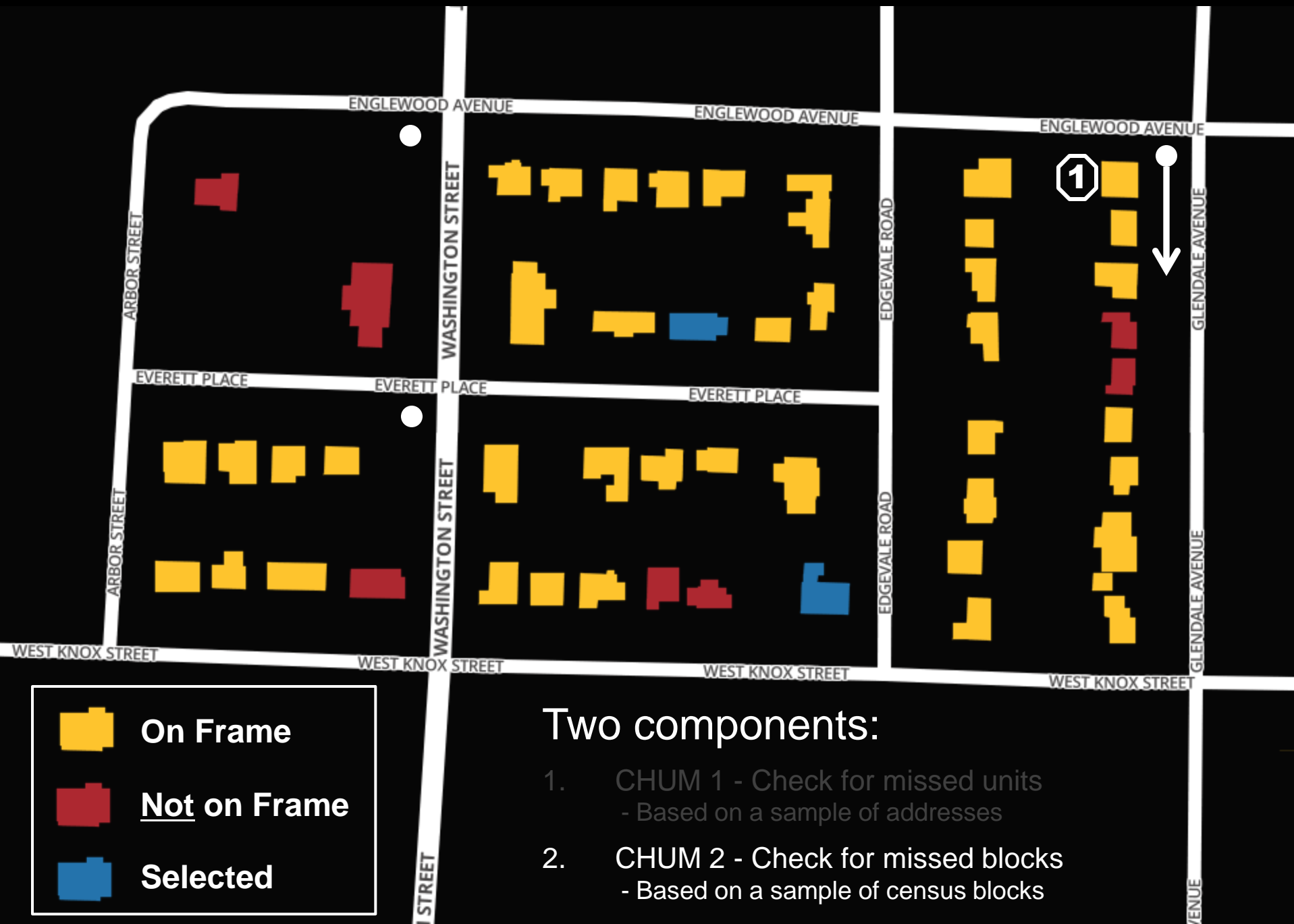
 On Frame




 Not on Frame

 Selected

Two components:

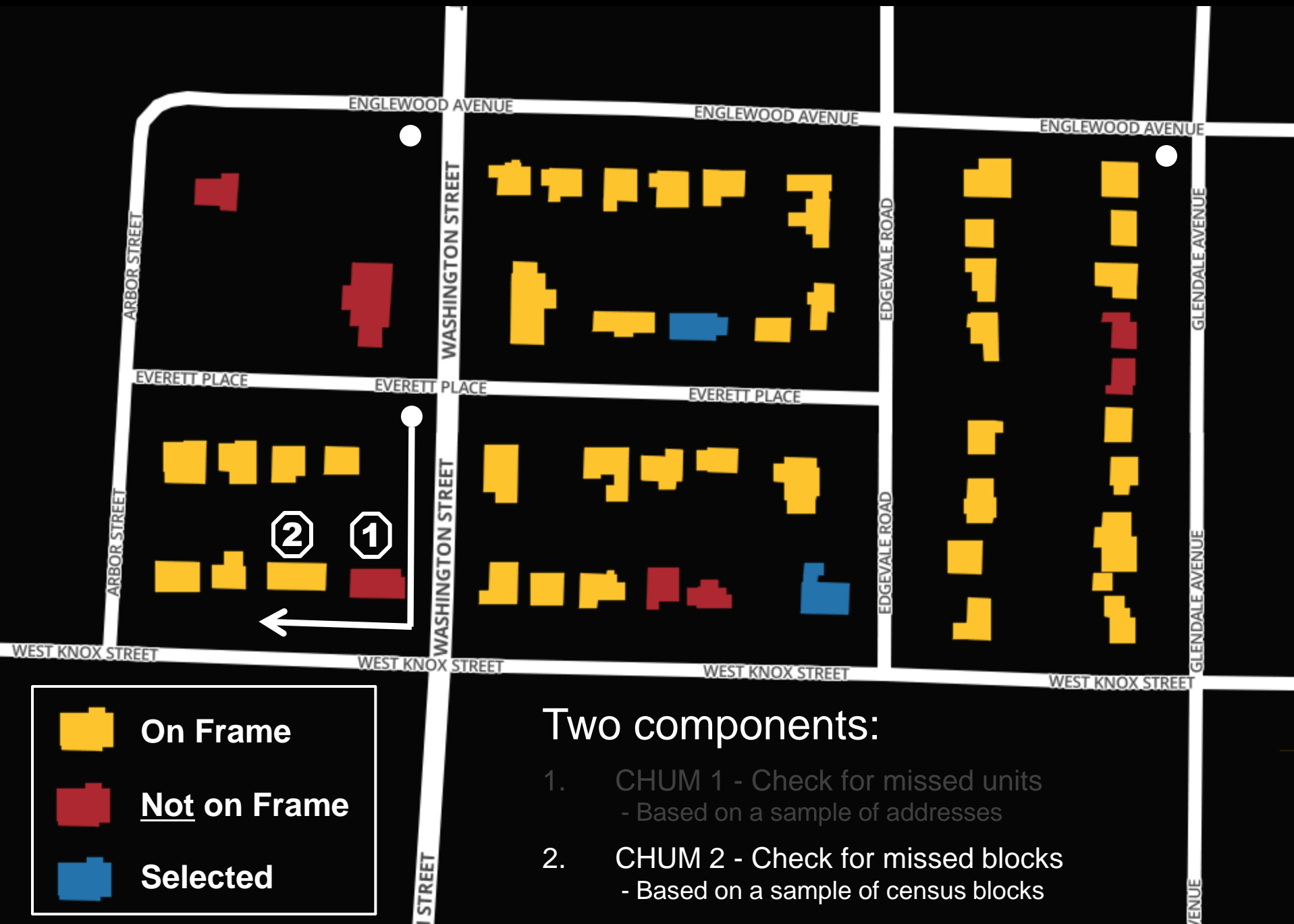
1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks



-  On Frame
-  Not on Frame
-  Selected

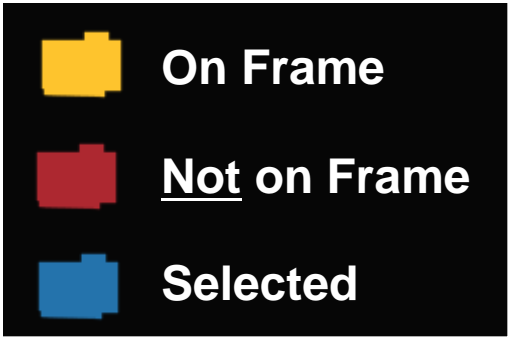
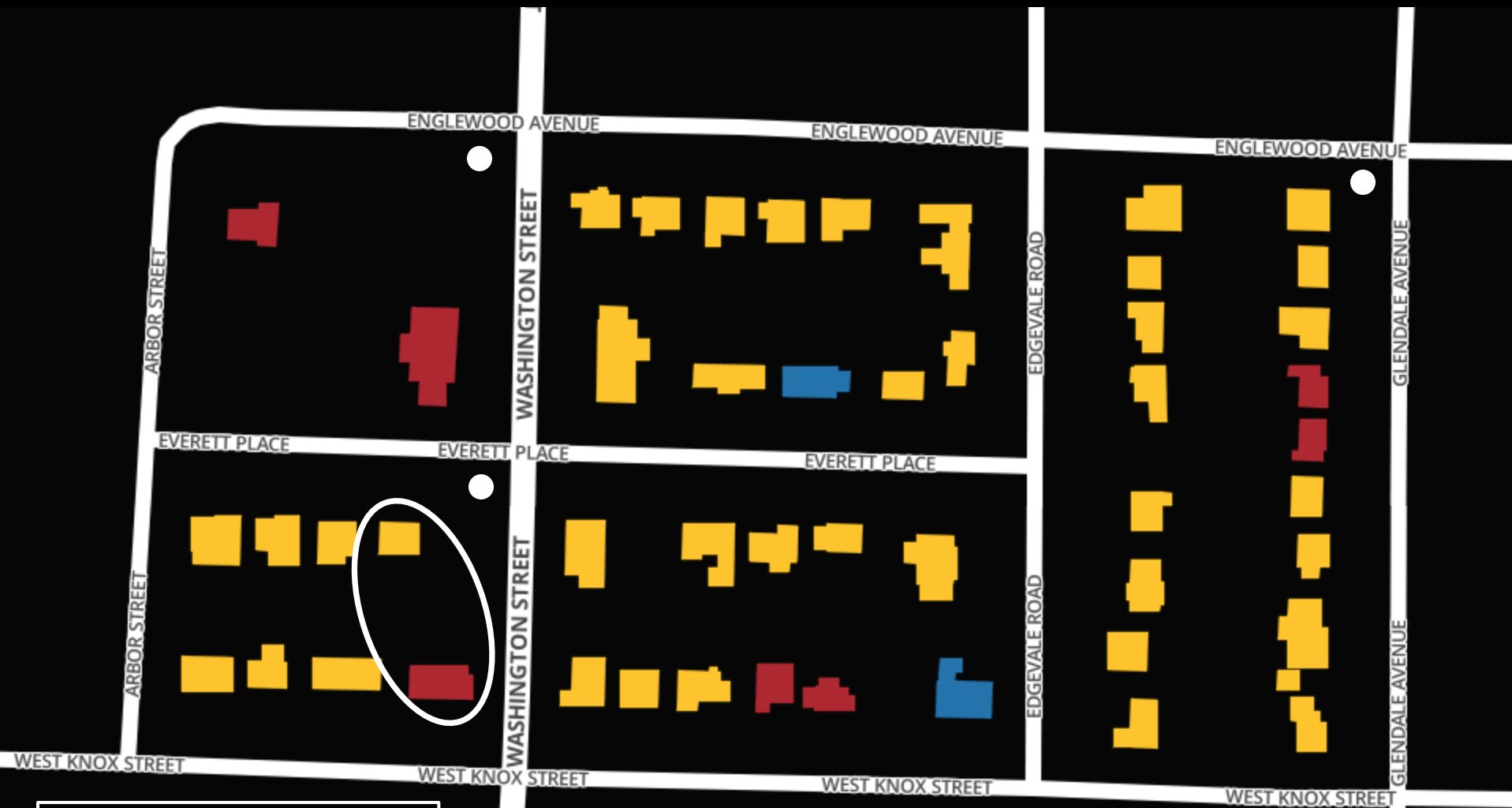
Two components:

1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks



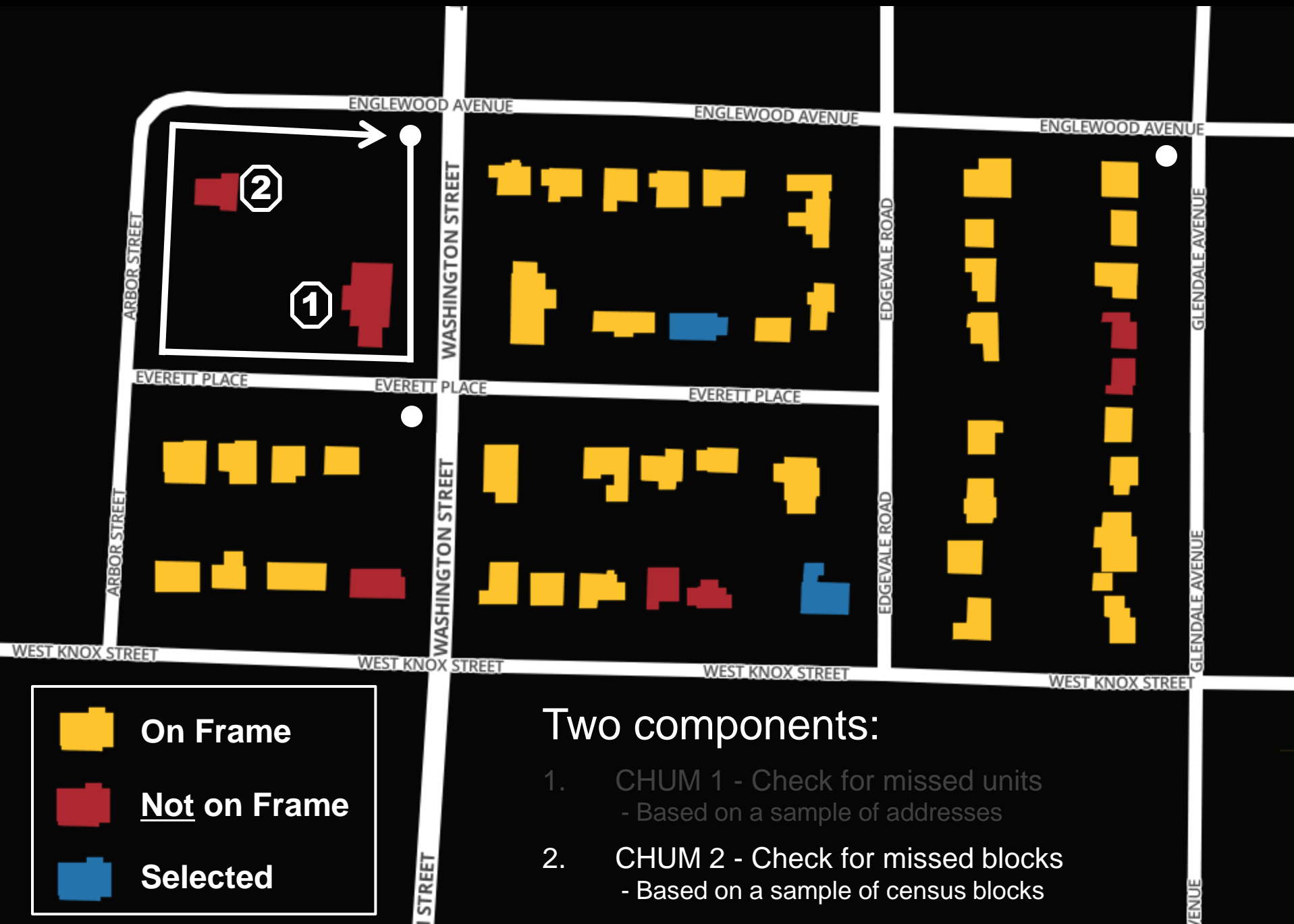
Two components:




1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks



Two components:

1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks



-  On Frame
-  Not on Frame
-  Selected

Two components:

1. CHUM 1 - Check for missed units
- Based on a sample of addresses
2. CHUM 2 - Check for missed blocks
- Based on a sample of census blocks

CHUM Summary

- For in-person surveys the CHUM systematically provides every dwelling unit, not on the ABS frame, a known probability of selection.
 - This is conceptually similar the half-open interval (HOI) procedure.
- CHUM procedure is applied to a sample.
- CHUM is typically implemented after sample selection.
- Two Components:
 1. CHUM 1 - Check for missed units.
 - Based on a sample of addresses
 2. CHUM 2 - Check for missed blocks.
 - Based on a sample of census blocks

Operational Considerations

Procedures need to be implemented correctly

- Training, monitoring and support are keys to successful implementation. Attempts to limit human error found in all field work.
- Technology is allowing for better implementation and monitoring. (handheld tablets, GPS, searchable database, electronic maps)

Keys to successful implementation

- Training
 - Classroom style
 - Video examples
 - Exercises
- Quality monitoring
 - Seeding (remove surrounding addresses from FI list so a miss unit must be found)
 - GPS (validates correct location and path of travel)
 - Site visits (when appropriate)
- Support for field staff
 - phone
 - email
 - real-time data transfer when possible

Potential Drawbacks and Benefits

Potential Drawbacks:

- Increased design effects from sub-sampling is likely.
- Can be more challenging to control sample size.
- Sample design needs to take into account estimated coverage.

Potential Benefits:

- CHUM does not limit the segment size (spatial or population) because every dwelling unit does not need to be visited. This is in contrast to listing procedures which need to manage segment size to reduce costs.
- CHUM can be implemented during the data collection time period after sample selection. This is in contrast to listing procedures which are done prior to sample selection. CHUM does not require separate visits to the segment.

Note: Frame supplementation methods are just one tool that attempt to increase data quality for a given cost. Specifically they attempt to reduce bias introduced by the frame. More work is needed to understand the cost components of these methods and how they relate to controlling the bias and variability of estimates so survey practitioners can choose the methods appropriate for their survey.

Case Studies

#1 - NSDUH Mailing List Field Studies

#2 - 2008 American National Election Studies

#3 - National Children's Study, Los Angeles County

#1 - NSDUH Mailing List Field Studies

Goals:

- Develop CHUM training materials and methods
- Test field staff implementation
- Understand the problems in implementation
- Develop a plan for improving methods

#1 - NSDUH Mailing List Field Studies

Field Test:

- CHUM scenarios created by teams of survey researchers.
- Scenarios not representative of population.
- Scenarios intentionally made difficult to help identify weaknesses in implementation and training methods.
- 20 field staff trained and sent to the field.

#1 - NSDUH Mailing List Field Studies

Lessons Learned:

- Field staff can successfully implement the CHUM procedure in most cases.
- Field staff need some access to support to work out the most challenging scenarios.
- Training can be simplified. Original training was too detailed and attempted to teach too much.
- Field staff performed better on day two once they got the hang of it.
- Need to better train field staff on when seek help and not solve every problem themselves.

Case Studies #2 and #3

#2 - 2008 American National Election Studies

- First full implementation of CHUM
- Segment size: census block group (CBG)
- 20 minute training (not long enough)
- Lots of field support was needed

#3 - National Children's Study, Los Angeles County

- First implementation after redesigning training materials
- Segment size: Smaller than CBG
- 4 hour training
- Very little field support was needed

Note: As a cost saving measure, this study first experimented with the use of graduate students to implement the CHUM procedure. Through the use of seeding it was determined the students were not correctly performing the CHUM so professional field staff redid the work. The following table reflects the work of professional field staff.

CHUM Quality Evaluation

		Seeded Address	Seeds Found	
			#	%
National Election Survey	Total	437	348	79.6
	CHUM 1	320	231	72.2
	CHUM 2	117	117	100.0
Children's Study	Total	53	50	94.3
	CHUM 1	53	50	94.3
	CHUM 2	0	-	-

Future and Current Research

- Mitigate design effects incurred by the CHUM through sample design
- Improve CHUM2 selection methods